```
seq_name: gb_ba2:AE000213

seq_documentation_block:
LOCUS       AE000213    10959 bp    DNA             BCT       12-NOV-1998
DEFINITION  Escherichia coli K-12 MG1655 section 103 of      f the complete
            genome.
ACCESSION   AE000213 U00096
VERSION     AE000213.1  GI:1787371
KEYWORDS    .
SOURCE      Escherichia coli.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 10959)
  AUTHORS   Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V.,
            Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F.,
            Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J.,
            Mau,B. and Shao,Y.
  TITLE     The complete genome sequence of Escherichia coli K-12
  JOURNAL   Science 277 (5331), 1453-1474 (1997)
  MEDLINE   97426617
REFERENCE   2  (bases 1 to 10959)
  AUTHORS   Blattner,F.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
            University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
            Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax:
            608-263-7459
REFERENCE   3  (bases 1 to 10959)
  AUTHORS   Blattner,F.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
            University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
            Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax:
            608-263-7459
REFERENCE   4  (bases 1 to 10959)
  AUTHORS   Plunkett,G. III.
  TITLE     Direct Submission

alignment_scores:
            Quality:  983.00              Length:    218
              Ratio:  4.795               Gaps:        0
  Percent Similarity:  94.037   Percent Identity: 85.780

alignment_block:
 US-09-252-691-7056 x AE000213/rev   ..

 Align seq 1/1  to reverse of: AE000213  from: 1  to: 10959

    1 AlaIleMetArgGlnLeuIleThrProGluAsnThrMetThrLysThrSe 17
      |||||||||||||||:::|||      ||||||||||||   ||||||||
 9268 GCTATAATGCGGCAATTCATAATCTCTGAAAATACCATGCAAAAAACTTC 9219

   17 rPheArgLysHisArgValGluArgPheSerSerArgGlnAlaThrArgA 34
      |||||||:::|||:::|||:::|||||||||||||:::::::::|||||||
 9218 TTTTAGAAATCACCAGGTTAAGCGATTCAGCTCGCAACGTTCTACCAGGC 9169

   34 rgThrProGluProGlnProThrArgValIleLeuPheAsnLysProTyr 50
      ||   ||||||   ||||||||||||||||||||||||||||||||||||
 9168 GTAAACCTGAAAACCAGCCCACGCGTGTGATCCTGTTCAATAAACCCTAC 9119

   51 AspValLeuProGlnPheThrAspGluAlaGlyArgSerThrLeuLysAs 67
      |||||||||||||||||||||||||||||||||:::|||||||||::
 9118 GATGTTCTTCCGCAGTTCACCGATGAAGCCGGACGCAAAACATTAAAAGA 9069

   67 pPheIleProValGlnGlyValTyrAlaAlaGlyArgLeuAspArgAspS 84
      :||||||||||||||||||||||||||||||||||||||||||||||||||
 9068 ATTCATCCCGGTTCAGGGTGTTTATGCAGCAGGTCGCCTTGACCGCGATA 9019

   84 erGluGlyLeuLeuValLeuThrAsnAspGlyValLeuGlnAlaArgLeu 100
      ||||||||||||||||||||||||||:::|||:::|||||||||||||||
 9018 GCGAAGGGTTGCTGGTGCTGACCAATAACGGCGCGTTGCAGGCGCGTTTA 8969

  101 ThrGlnProGlyLysArgThrGlyLysIleTyrTyrValGlnValGluGl 117
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 8968 ACCCAGCCGGGTAAACGCACCGGAAAAATCTATTATGTGCAGGTGGAAGG 8919

  117 yGluProAspAspAlaSerLeuAlaLysLeuArgAsnGlyValThrLeuA 134
      |   |||   :::   :::|||      ||||||||||||||||||||||
 8918 TATTCCCACACAAGACGCACTTGAAGCGTTGCGCAATGGCGTAACCTTAA 8869

  134 snAspGlyProThrLeuProAlaGlyIleGluArgValAsnGluProGlu 150
      ||||||||||||||||||||||||||  |||   |||:::||||||
 8868 ATGATGGCCCTACCCTGCCCGCCGGCGCGGAGCTGGTTGACGAACCCGCG 8819

  151 TrpLeuTrpProArgAsnProProIleArgGluArgLysSerIleProTh 167
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 8818 TGGTTATGGCCGCGGAATCCACCAATTCGTGAACGCAAAAGTATTCCCAC 8769

  167 rSerTrpLeuLysIleThrLeuTyrGluGlyArgAsnArgGlnValArgA 184
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 8768 CAGCTGGCTGAAGATCACCTTATATGAAGGACGTAATCGCCAGGTGCGCC 8719

  184 rgMetThrAlaHisValGlyProProThrLeuArgLeuIleArgTyrAla 200
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 8718 GCATGACCGCCCATGTTGGCTTCCCCACGCTGCGACTGATTCGCTATGCG 8669

  201 MetGlySerTyrThrLeuAspSerLeuAlaAsnGlyGluTrpArgAspVa 217
      ||||||:::|||:::|||||:::||||||||||||||||||||||:::||
 8668 ATGGGTGATTACTCTTTGGATAATCTTGCCAATGGTGAATGGCGAGAAGT 8619

  217 lThr 218
      ||||
 8618 GACA 8615
```

ALIGNMENTS

```
RESULT   1
YMFC_ECOLI
ID   YMFC_ECOLI      STANDARD;      PRT;   217 AA.
AC   P75966;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-FEB-2000 (Rel. 39, Last annotation update)
DE   HYPOTHETICAL 24.9 KD PROTEIN IN TRMU-ICDA INTERGENIC REGION.
GN   YMFC.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE; 97426617.
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE; 97061202.
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF PSEUDOURIDINE SYNTHASES.
CC       STRONG, TO H.INFLUENZAE HI0694.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE000213; AAC74219.1; ALT_INIT.
DR   EMBL; D90748; BAA35957.1; -.
DR   EMBL; D90749; BAA35966.1; -.
DR   ECOGENE; EG13447; YMFC.
DR   PROSITE; PS01149; PSI_RSU; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   217 AA;  24880 MW;  F7C7A7CEDC5FD3F6 CRC64;

  Query Match             83.7%;  Score 975;  DB 1;  Length 217;
  Best Local Similarity   85.6%;  Pred. No. 3.8e-79;

  Matches  185;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       3 MRQLITPENTMTKTSFRKHRVERFSSRQATRRTPEPQPTRVILFNKPYDVLPQFTDEAGR 62
           |||  |  |||| ||||| |:|\:||||:::||| || ||||||||||||||||||||||||
Db       1 MRQFIISENTMQKTSFRNHQVKRFSSQRSTRRKPENQPTRVILFNKPYDVLPQFTDEAGR 60

Qy      63 STLKDFIPVQGVYAAGRLDRDSEGLLVLTNDGVLQARLTQPGKRTGKIYYVQVEGEPDDA 122
           |||:|||||||||||||||||||||||:| |||||||||||||||||||| |
Db      61 KTLKEFIPVQGVYAAGRLDRDSEGLLVLTNNGALQARLTQPGKRTGKIYYVQVEGIPTQD 120

Qy     123 SLAKLRNGVTLNDGPTLPAGIERVNEPEWLWPRNPPIRERKSIPTSWLKITLYEGRNRQV 182
           :|  |||||||||||||||| | |:|| ||||||||||||||||||||||||||||||||
Db     121 ALEALRNGVTLNDGPTLPAGAELVDEPAWLWPRNPPIRERKSIPTSWLKITLYEGRNRQV 180

Qy     183 RRMTAHVGFPTLRLIRYAMGSYTLDSLANGEWRDVT 218
           |||||||||||||||||| |:||:|||||||:||
Db     181 RRMTAHVGFPTLRLIRYAMGDYSLDNLANGEWREVT 216
```

RESULT 10
W10941
ID W10941 standard; peptide; 35 AA.
AC W10941;
DT 10-NOV-1997 (first entry)
DE Polyclonal anti-ferritin binder sequence, C28, from R26 library.
KW Functional surrogate; analyte; affinity receptor; immunoreactive group;
KW mimic; homogenous immunoassay; detection; diagnostic analyte; Chlamydia;
KW cardiac marker; tumour marker; allergen; hormone; fertility; myoglobin;
KW pregnancy; infectious disease; ferritin; myosin light chain; troponin;
KW follicle stimulating hormone; human; growth hormone; immunoglobulin E;
KW prolactin; parathyroid hormone; placental lactogen; hepatitis antigen;
KW antibody; chorionic gonadotropin; luteinising hormone; cytomegalovirus;
KW Streptococcus; rubella; toxoplasma; DK-MB; prostate-specific antigen;
KW carcinoembryonic antigen; alpha-fetoprotein; herpes virus; CA125.
OS Synthetic.
PN WO9641172-A1.
PD 19-DEC-1996.
PF 07-JUN-1996; U10498.
PR 07-JUN-1995; US-476375.
PA (CYTO-) CYTOGEN CORP.
PI Carter JM, Lee-Own FV;
DR WPI; 97-077284/07.
DR N-PSDB; T48787.
PT Labelled functional surrogate of an analyte - useful as competitor
PT molecule in affinity assays, esp. for detecting large macromolecules
PT such as ferritin
PS Disclosure; Page 118/2; 156pp; English.
CC This sequence represents a polyclonal anti-ferritin binder sequence from
CC the R26 library (C series) which may be used in the conjugate of the

CC invention. The novel labelled conjugate comprises at least one label
CC attached to a functional surrogate of an analyte of interest. The
CC surrogate is capable of competing effectively with the analyte for a
CC limiting amount of an affinity receptor for the analyte. The conjugate
CC exhibits an activity that is altered upon interaction with the affinity
CC receptor and this activity can be measured and related to the amount of
CC the analyte present in a sample. Functional surrogates such as this have
CC an immunoreactive group that allows the surrogate to compete effectively
CC and with the analyte for a limiting amount of its affinity receptor.
CC Functional surrogates are able to mimic naturally occurring analytes.
CC They can be labelled for use in standard competitive affinity assays
CC (esp. homogenous immunoassays) for detecting large macromolecules such
CC as polypeptides, polysaccharides, polynucleotides, glycoproteins and
CC lipid-containing macromolecules, as well as small haptens. Typical
CC diagnostic analytes for detection include cardiac or tumour markers,
CC allergens, hormones related to fertility-pregnancy or analytes associated
CC with infectious disease. In particular, the assays are useful for
CC detecting ferritin, follicle stimulating hormone, human growth hormone,
CC immunoglobulin E, prolactin, parathyroid hormone, human placental
CC lactogen, hepatitis antigens or antibodies against them, human
CC chorionic gonadotropin, human luteinising hormone, cytomegalovirus,
CC Chlamydia, Streptococcus a, rubella, toxoplasma, herpes virus, DK-MB,
CC myoglobin, myosin light chain, troponin, carcinoembryonic antigen,
CC alpha-fetoprotein, prostate-specific antigen and CA125 (a tumour marker).
SQ Sequence 35 AA;

Query Match 2.7%; Score 6; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 AAGRLD 81
||||||
Db 17 AAGRLD 22